



ENTERED

OIPE

RAW SEQUENCE LISTING

DATE: 02/04/2003 P.6

PATENT APPLICATION: US/09/988,067B

TIME: 14:28:29

Input Set : A:\06132\040002\SEQLIST.TXT

Output Set: N:\CRF4\02042003\I988067B.raw

4 <110> APPLICANT: Haas, Rainer
 5 Kleanthous, Harold
 6 Tomb, Jean-Francois
 7 Miller, Charles
 8 Al-Garawi, Amal
 9 Odenbreit, Stefan
 10 Meyer, Thomas
 14 <120> TITLE OF INVENTION: Helicobacter Polypeptides and
 15 Corresponding Polynucleotide Molecules
 18 <130> FILE REFERENCE: 06132/040002
 20 <140> CURRENT APPLICATION NUMBER: US 09/988,067B
 C--> 21 <141> CURRENT FILING DATE: 2003-01-31
 23 <150> PRIOR APPLICATION NUMBER: US 08/831,309
 24 <151> PRIOR FILING DATE: 1997-04-01
 26 <160> NUMBER OF SEQ ID NOS: 112
 28 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 30 <210> SEQ ID NO: 1
 31 <211> LENGTH: 989
 32 <212> TYPE: DNA
 33 <213> ORGANISM: Helicobacter pylori
 35 <220> FEATURE:
 36 <221> NAME/KEY: CDS
 37 <222> LOCATION: (71)...(940)
 39 <400> SEQUENCE: 1
 40 ctatgacgat tgtctcgctt ttagaaaaca ctctaatacgc ttttgaaaaa caacaaagga 60
 41 agggatttta atg aaa ttt tta cgc tct gtt tat gca ttt tgc tcc agt 109
 42 Met Lys Phe Leu Arg Ser Val Tyr Ala Phe Cys Ser Ser
 43 1 5 10
 45 tgg gta ggg acg att gtt att gtg ctg ttg gtt atc ttt ttt atc gcg 157
 46 Trp Val Gly Thr Ile Val Ile Val Leu Leu Val Ile Phe Phe Ile Ala
 47 15 20 25
 49 caa gcc ttt atc att ccc tct cgc tct atg gtt ggc acg ctc tat gag 205
 50 Gln Ala Phe Ile Ile Pro Ser Arg Ser Met Val Gly Thr Leu Tyr Glu
 51 30 35 40 45
 53 ggc gac atg ctc ttt gtc aaa aag ttt tct tac ggc ata ccc att cct 253
 54 Gly Asp Met Leu Phe Val Lys Lys Phe Ser Tyr Gly Ile Pro Ile Pro
 55 50 55 60
 57 aaa atc cca tgg att gag ctt cct gtt atg cct gat ttt aaa aat aac 301
 58 Lys Ile Pro Trp Ile Glu Leu Pro Val Met Pro Asp Phe Lys Asn Asn
 59 65 70 75
 61 gga cat ttg ata gag ggg gat cgc cct aag cgt ggc gaa gtg gtg gtg 349
 62 Gly His Leu Ile Glu Gly Asp Arg Pro Lys Arg Gly Glu Val Val Val
 63 80 85 90

RAW SEQUENCE LISTING

DATE: 02/04/2003

PATENT APPLICATION: US/09/988,067B

TIME: 14:28:29

Input Set : A:\06132.040002.SEQLIST.TXT

Output Set: N:\CRF4\02042003\I988067B.raw

```

65 ttt atc cct ccc cat gaa aaa aag tct tac tat gtt aaa agg aat ttt 397
66 Phe Ile Pro Pro His Glu Lys Lys Ser Tyr Tyr Val Lys Arg Asn Phe
67 95 100 105
69 gcc att gga ggc gat gag gtg ttg ttc act aat gag ggt ttt tat ttg 445
70 Ala Ile Gly Gly Asp Glu Val Leu Phe Thr Asn Glu Gly Phe Tyr Leu
71 110 115 120 125
73 cac cct ttt gag agc gac acg gac aaa aat tac atc gct aaa cat tac 493
74 His Pro Phe Glu Ser Asp Thr Asp Lys Asn Tyr Ile Ala Lys His Tyr
75 130 135 140
77 cct aac gcc atg aca aaa gaa ttt atg ggt aaa att ttt gtt tta aac 541
78 Pro Asn Ala Met Thr Lys Glu Phe Met Gly Lys Ile Phe Val Leu Asn
79 145 150 155
81 cct tat aaa aat gag cat ccg ggt atc cat tac caa aaa gac aat gaa 589
82 Pro Tyr Lys Asn Glu His Pro Gly Ile His Tyr Gln Lys Asp Asn Glu
83 160 165 170
85 acc ttc cac tta atg gag caa tta gcc act caa ggc gca gaa gct aat 637
86 Thr Phe His Leu Met Glu Gln Leu Ala Thr Gln Gly Ala Glu Ala Asn
87 175 180 185
89 atc agc atg caa ctc att caa atg gag ggc gaa aag gtg ttt tat aag 685
90 Ile Ser Met Gln Leu Ile Gln Met Glu Gly Glu Lys Val Phe Tyr Lys
91 190 195 200 205
93 aaa atc aat gac gat gaa ttt ttc atg atc ggc gac aac aga gac aat 733
94 Lys Ile Asn Asp Asp Glu Phe Phe Met Ile Gly Asp Asn Arg Asp Asn
95 210 215 220
97 tct agc gac tcg cgc ttt tgg ggg agt gtg gct tat aaa aac atc gtg 781
98 Ser Ser Asp Ser Arg Phe Trp Gly Ser Val Ala Tyr Lys Asn Ile Val
99 225 230 235
101 ggt tcg cca tgg ttt gtt tat ttc agt ttg agt tta aaa aat agc cta 829
102 Gly Ser Pro Trp Phe Val Tyr Phe Ser Leu Ser Leu Lys Asn Ser Leu
103 240 245 250
105 gaa atg gat gca gaa aat aac cct aaa aaa cgc tat ctg gtg cgt tgg 877
106 Glu Met Asp Ala Glu Asn Asn Pro Lys Lys Arg Tyr Leu Val Arg Trp
107 255 260 265
109 gaa cgc atg ttt aaa agc gtt gga ggc tta gaa aaa atc att aaa aaa 925
110 Glu Arg Met Phe Lys Ser Val Gly Gly Leu Glu Lys Ile Ile Lys Lys
111 270 275 280 285
113 gaa aac gca acg cat taagggtttt tgtgcaattt tttgatttct ctttagaaaag 980
114 Glu Asn Ala Thr His
115 290
117 ttttattac 989
119 <210> SEQ ID NO: 2
120 <211> LENGTH: 290
121 <212> TYPE: PRT
122 <213> ORGANISM: Helicobacter pylori
124 <400> SEQUENCE: 2
125 Met Lys Phe Leu Arg Ser Val Tyr Ala Phe Cys Ser Ser Trp Val Gly
126 1 5 10 15
127 Thr Ile Val Ile Val Leu Leu Val Ile Phe Phe Ile Ala Gln Ala Phe
128 20 25 30

```

RAW SEQUENCE LISTING

DATE: 02/04/2003

PATENT APPLICATION: US/09/988,067B

TIME: 14:28:29

Input Set : A:\06132.040002.SEQLIST.TXT

Output Set: N:\CRF4\02042003\I988067B.raw

```

129 Ile Ile Pro Ser Arg Ser Met Val Gly Thr Leu Tyr Glu Gly Asp Met
130      35      40      45
131 Leu Phe Val Lys Lys Phe Ser Tyr Gly Ile Pro Ile Pro Lys Ile Pro
132      50      55      60
133 Trp Ile Glu Leu Pro Val Met Pro Asp Phe Lys Asn Asn Gly His Leu
134 65      70      75      80
135 Ile Glu Gly Asp Arg Pro Lys Arg Gly Glu Val Val Val Phe Ile Pro
136      85      90      95
137 Pro His Glu Lys Lys Ser Tyr Tyr Val Lys Arg Asn Phe Ala Ile Gly
138      100     105     110
139 Gly Asp Glu Val Leu Phe Thr Asn Glu Gly Phe Tyr Leu His Pro Phe
140      115     120     125
141 Glu Ser Asp Thr Asp Lys Asn Tyr Ile Ala Lys His Tyr Pro Asn Ala
142      130     135     140
143 Met Thr Lys Glu Phe Met Gly Lys Ile Phe Val Leu Asn Pro Tyr Lys
144 145      150     155     160
145 Asn Glu His Pro Gly Ile His Tyr Gln Lys Asp Asn Glu Thr Phe His
146      165     170     175
147 Leu Met Glu Gln Leu Ala Thr Gln Gly Ala Glu Ala Asn Ile Ser Met
148      180     185     190
149 Gln Leu Ile Gln Met Glu Gly Glu Lys Val Phe Tyr Lys Lys Ile Asn
150      195     200     205
151 Asp Asp Glu Phe Phe Met Ile Gly Asp Asn Arg Asp Asn Ser Ser Asp
152      210     215     220
153 Ser Arg Phe Trp Gly Ser Val Ala Tyr Lys Asn Ile Val Gly Ser Pro
154 225      230     235     240
155 Trp Phe Val Tyr Phe Ser Leu Ser Leu Lys Asn Ser Leu Glu Met Asp
156      245     250     255
157 Ala Glu Asn Asn Pro Lys Lys Arg Tyr Leu Val Arg Trp Glu Arg Met
158      260     265     270
159 Phe Lys Ser Val Gly Gly Leu Glu Lys Ile Ile Lys Lys Glu Asn Ala
160      275     280     285
161 Thr His
162      290
165 <210> SEQ ID NO: 3
166 <211> LENGTH: 514
167 <212> TYPE: DNA
168 <213> ORGANISM: Helicobacter pylori
170 <220> FEATURE:
171 <221> NAME/KEY: CDS
172 <222> LOCATION: (112)...(471)
174 <400> SEQUENCE: 3
175 ggatttttta gagctcttag tcaatgataa tgtggtagaa acgattgaaa aaggctttgt 60
176 gatagggtttt ggagcggggg atattaccta tcaattaaga ggcgaaatgt a atg ggt 117
177      Met Gly
178      1
180 gca gtg gtt gtt tta ttt tta acg ctg gtt tta ttg ttt tta gtt tta 165
181 Ala Val Val Val Leu Phe Leu Thr Leu Val Leu Leu Phe Leu Val Leu
182      5      10      15

```

RAW SEQUENCE LISTING

DATE: 02/04/2003

PATENT APPLICATION: US/09/988,067B

TIME: 14:28:29

Input Set : A:\06132.040002.SEQLIST.TXT

Output Set: N:\CRF4\02042003\I988067B.raw

```

184 agg gat ttt ggt tta gca agc ccc aaa caa aag att tta gct ttt tta 213
185 Arg Asp Phe Gly Leu Ala Ser Pro Lys Gln Lys Ile Leu Ala Phe Leu
186      20      25      30
188 atc gta ggg att ata gga gcg agc atc agc gtt tat act tac aag caa 261
189 Ile Val Gly Ile Ile Gly Ala Ser Ile Ser Val Tyr Thr Tyr Lys Gln
190 35      40      45      50
192 aac caa caa aac caa caa gag atc gct ttg caa aga gcg ttt tta agg 309
193 Asn Gln Gln Asn Gln Gln Glu Ile Ala Leu Gln Arg Ala Phe Leu Arg
194      55      60      65
196 ggg gaa acc ttg ttg tgt aaa ggc att aaa gtc aat aac caa acc ttt 357
197 Gly Glu Thr Leu Leu Cys Lys Gly Ile Lys Val Asn Asn Gln Thr Phe
198      70      75      80
200 aat tta gtg agc gga act tta agc ttt tta ggc aaa aaa caa acc cct 405
201 Asn Leu Val Ser Gly Thr Leu Ser Phe Leu Gly Lys Lys Gln Thr Pro
202      85      90      95
204 atg aaa gac gtt ctt gtg gat ttg gat tct tgt cag acg ctc caa aaa 453
205 Met Lys Asp Val Leu Val Asp Leu Asp Ser Cys Gln Thr Leu Gln Lys
206      100      105      110
208 gat ccc tta atc caa ccc taatgatgaa taataataat accccaccca 501
209 Asp Pro Leu Ile Gln Pro
210 115      120
212 aaccctaga aga 514
214 <210> SEQ ID NO: 4
215 <211> LENGTH: 120
216 <212> TYPE: PRT
217 <213> ORGANISM: Helicobacter pylori
219 <400> SEQUENCE: 4
220 Met Gly Ala Val Val Val Leu Phe Leu Thr Leu Val Leu Leu Phe Leu
221 1      5      10      15
222 Val Leu Arg Asp Phe Gly Leu Ala Ser Pro Lys Gln Lys Ile Leu Ala
223      20      25      30
224 Phe Leu Ile Val Gly Ile Ile Gly Ala Ser Ile Ser Val Tyr Thr Tyr
225      35      40      45
226 Lys Gln Asn Gln Gln Asn Gln Gln Glu Ile Ala Leu Gln Arg Ala Phe
227      50      55      60
228 Leu Arg Gly Glu Thr Leu Leu Cys Lys Gly Ile Lys Val Asn Asn Gln
229 65      70      75      80
230 Thr Phe Asn Leu Val Ser Gly Thr Leu Ser Phe Leu Gly Lys Lys Gln
231      85      90      95
232 Thr Pro Met Lys Asp Val Leu Val Asp Leu Asp Ser Cys Gln Thr Leu
233      100      105      110
234 Gln Lys Asp Pro Leu Ile Gln Pro
235      115      120
238 <210> SEQ ID NO: 5
239 <211> LENGTH: 1233
240 <212> TYPE: DNA
241 <213> ORGANISM: Helicobacter pylori
243 <220> FEATURE:
244 <221> NAME/KEY: CDS

```

RAW SEQUENCE LISTING

DATE: 02/04/2003

PATENT APPLICATION: US/09/988,067B

TIME: 14:28:29

Input Set : A:\06132.040002.SEQLIST.TXT

Output Set: N:\CRF4\02042003\I988067B.raw

245 <222> LOCATION: (135)...(1049)

247 <400> SEQUENCE: 5

```

248 gtttttaatt taatattcat taagcttttg tggctattcc attttaattt tgtttttcat 60
249 taaaacccaa tctaaaatct tatttttatg ataaaaatacc taatcataat atcaaattctt 120
250 aaaccaacga aacc atg aaa aaa gct ctc tta cta act ctc tct ctc tcg 170
251 Met Lys Lys Ala Leu Leu Leu Thr Leu Ser Leu Ser
252 1 5 10
254 ttc tgg ctc cac gct gaa agg aat gga ttt tat tta ggt tta aat ttt 218
255 Phe Trp Leu His Ala Glu Arg Asn Gly Phe Tyr Leu Gly Leu Asn Phe
256 15 20 25
258 cta gaa gga agc tat att aaa gga caa ggt agc atc ggc aaa aaa gct 266
259 Leu Glu Gly Ser Tyr Ile Lys Gly Gln Gly Ser Ile Gly Lys Lys Ala
260 30 35 40
262 tca gca gaa aac gcc tta aat gaa gcg atc aat aac gca aaa aat tca 314
263 Ser Ala Glu Asn Ala Leu Asn Glu Ala Ile Asn Asn Ala Lys Asn Ser
264 45 50 55 60
266 tta ttc cct aac aca aaa gcc ata aga gat gca caa aac gcc tta aat 362
267 Leu Phe Pro Asn Thr Lys Ala Ile Arg Asp Ala Gln Asn Ala Leu Asn
268 65 70 75
270 gca gtg aaa gat tca aac aaa atc gct agc cga ttc gca gga aat ggt 410
271 Ala Val Lys Asp Ser Asn Lys Ile Ala Ser Arg Phe Ala Gly Asn Gly
272 80 85 90
274 gga tcg ggc ggt ctt ttt aat gag ctc agc ttt ggg tat aaa tat ttt 458
275 Gly Ser Gly Gly Leu Phe Asn Glu Leu Ser Phe Gly Tyr Lys Tyr Phe
276 95 100 105
278 ttg ggt aaa aaa agg att ata ggg ttt agg cac tct ctt ttt ttc ggt 506
279 Leu Gly Lys Lys Arg Ile Ile Gly Phe Arg His Ser Leu Phe Phe Gly
280 110 115 120
282 tac caa ctt ggt ggc gtt ggt tct gtt cct ggt agc ggt tta atc gtt 554
283 Tyr Gln Leu Gly Gly Val Gly Ser Val Pro Gly Ser Gly Leu Ile Val
284 125 130 135 140
286 ttt tta ccc tat ggt ttc aat acg gat ttg ctc att aat tgg act aac 602
287 Phe Leu Pro Tyr Gly Phe Asn Thr Asp Leu Leu Ile Asn Trp Thr Asn
288 145 150 155
290 gat aag cga gcg tcc caa aaa tat gtt gaa cga agg gta aaa ggg ctc 650
291 Asp Lys Arg Ala Ser Gln Lys Tyr Val Glu Arg Arg Val Lys Gly Leu
292 160 165 170
294 tct ata ttt tac aaa gat atg acc ggc aga acg cta gac gct aat aca 698
295 Ser Ile Phe Tyr Lys Asp Met Thr Gly Arg Thr Leu Asp Ala Asn Thr
296 175 180 185
298 tta aaa aaa gca tca agg cat gta ttt aga aaa tct tca ggg ctt gtg 746
299 Leu Lys Lys Ala Ser Arg His Val Phe Arg Lys Ser Ser Gly Leu Val
300 190 195 200
302 att ggc atg gaa cta ggg ggt agc act tgg ttt gca agt aac aat ctc 794
303 Ile Gly Met Glu Leu Gly Gly Ser Thr Trp Phe Ala Ser Asn Asn Leu
304 205 210 215 220
306 acc cct ttc aat caa gtc aag agt cgc acg att ttt cag ttg caa gga 842
307 Thr Pro Phe Asn Gln Val Lys Ser Arg Thr Ile Phe Gln Leu Gln Gly
308 225 230 235

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/988,067B

DATE: 02/04/2003
TIME: 14:28:30

Input Set : A:\06132.040002.SEQLIST.TXT
Output Set: N:\CRF4\02042003\I988067B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:33; Xaa Pos. 176

Seq#:34; Xaa Pos. 176

Seq#:71; Xaa Pos. 496

Seq#:72; Xaa Pos. 496

VERIFICATION SUMMARY

DATE: 02/04/2003

PATENT APPLICATION: US/09/988,067B

TIME: 14:28:30

Input Set : A:\06132.040002.SEQLIST.TXT

Output Set: N:\CRF4\02042003\I988067B.raw

L:21 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:2483 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ ID#:33
L:2483 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:33 after pos.:584
L:2571 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:34 after pos.:160
L:5250 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ ID#:71
L:5250 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:71 after pos.:1602
L:5345 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:72 after pos.:480